Welcome

NOAA Comprehensive Large Array-data Stewardship System (CLASS)

Alexander Kidd /CLASS Project Area Lead

August 2005 Users' Workshop

Agenda

- **CLASS Mission**
- **CLASS Background**
- **CLASS Goals**
- **CLASS Overview**
- **CLASS Organization**
- **Accomplishments**
- **System Overview**
- **Hardware and Network Design**
- **Requirements Affecting Hardware**
- **Multi-Site Configuration**
- **Hardware Design Features**
- **CLASS Single Site Configuration**
- **Hardware Performance**

CLASS Mission

NOAA's National Data Centers and their world-wide clientele of customers look to CLASS as the sole NOAA IT infrastructure project in which all NOAA's current and future environmental data sets will reside. CLASS provides permanent, secure storage, and safe, efficient data discovery and access between the Data Centers and the customers.

CLASS Background

•The CLASS project derives in part from an effort by NOAA to centralize its numerous systems for (satellite) data access.

•The goal of this effort is to eliminate the various "stove-pipe" systems and produce a unified "enterprise" access system for the NOAA environmental data holdings.

CLASS Goals

- Give any potential customer access to all NOAA and non-NOAA data through a single portal.

- Eliminate the need to keep creating "stovepipe" systems for each new type of data, but, in as much as possible, use already polished portions/modules of existing legacy systems.

- Describe a cost-effective architecture that can primarily handle large-array data sets, with the capability of handling smaller ones as well.

CLASS Overview

- CLASS is a web-based data archive and distribution system for NOAA/NESDIS environmental data
- CLASS is an extension of an 1995 operational system … SAA (Satellite Active Archive)
  - Transition to the CLASS architecture began in 2001
  - CLASS became the Operational Archive and Access System for NOAA in 2004
- CLASS currently supports POES, DMSP, and GOES data sets
- CLASS will support additional campaigns, broader user base, new functionality as it evolves
  - CLASS concurrently supports ongoing operations and new requirements implementation

CLASS Overview
(Continued)

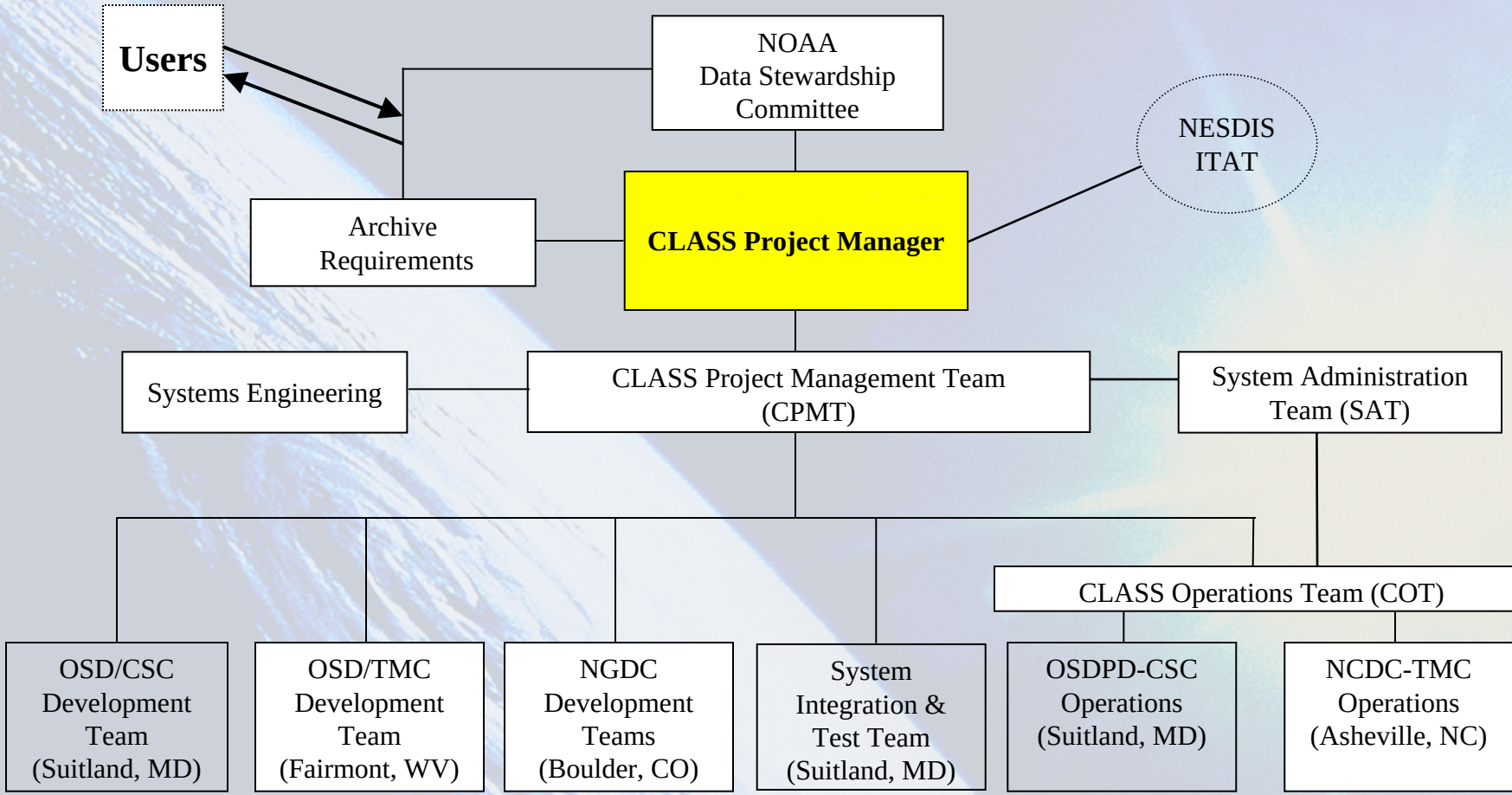
- Provide one stop shopping and access capability for NOAA and NESDIS environmental data and products
- Provide a common look and feel for accessing NOAA and NESDIS environmental data and products
- Provide an efficient architecture for archiving and distribution of NOAA and NESDIS environmental data and products
- Reduce implementation costs by using reengineering, and evolutionary effort
- Allow NOAA to fulfill its requirements regarding archive, access and distribution of large array data sets

CLASS Overview

(Continued)

- Accommodate expanding number of data sources
  - MetOp, NPP, NPOESS, EOS, In-situ, NexRAD, GOES-R, etc.
- Data volume is growing exponentially
  - Anticipating up to 100 Petabytes by 2015
- User volume is also growing exponentially
  - 1995 11,000 data sets delivered
  - 2004 4,216,000 data sets delivered

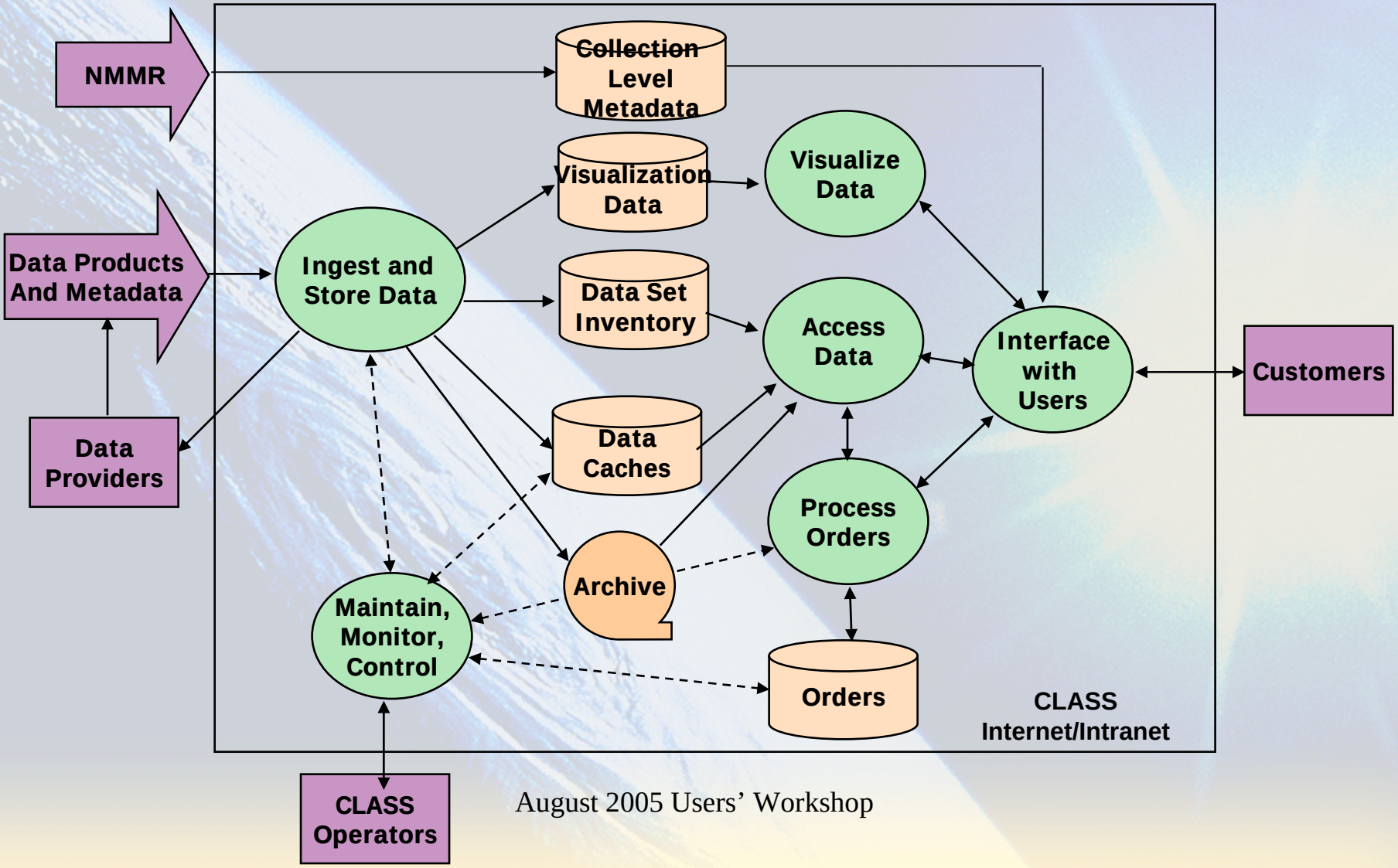
Organization



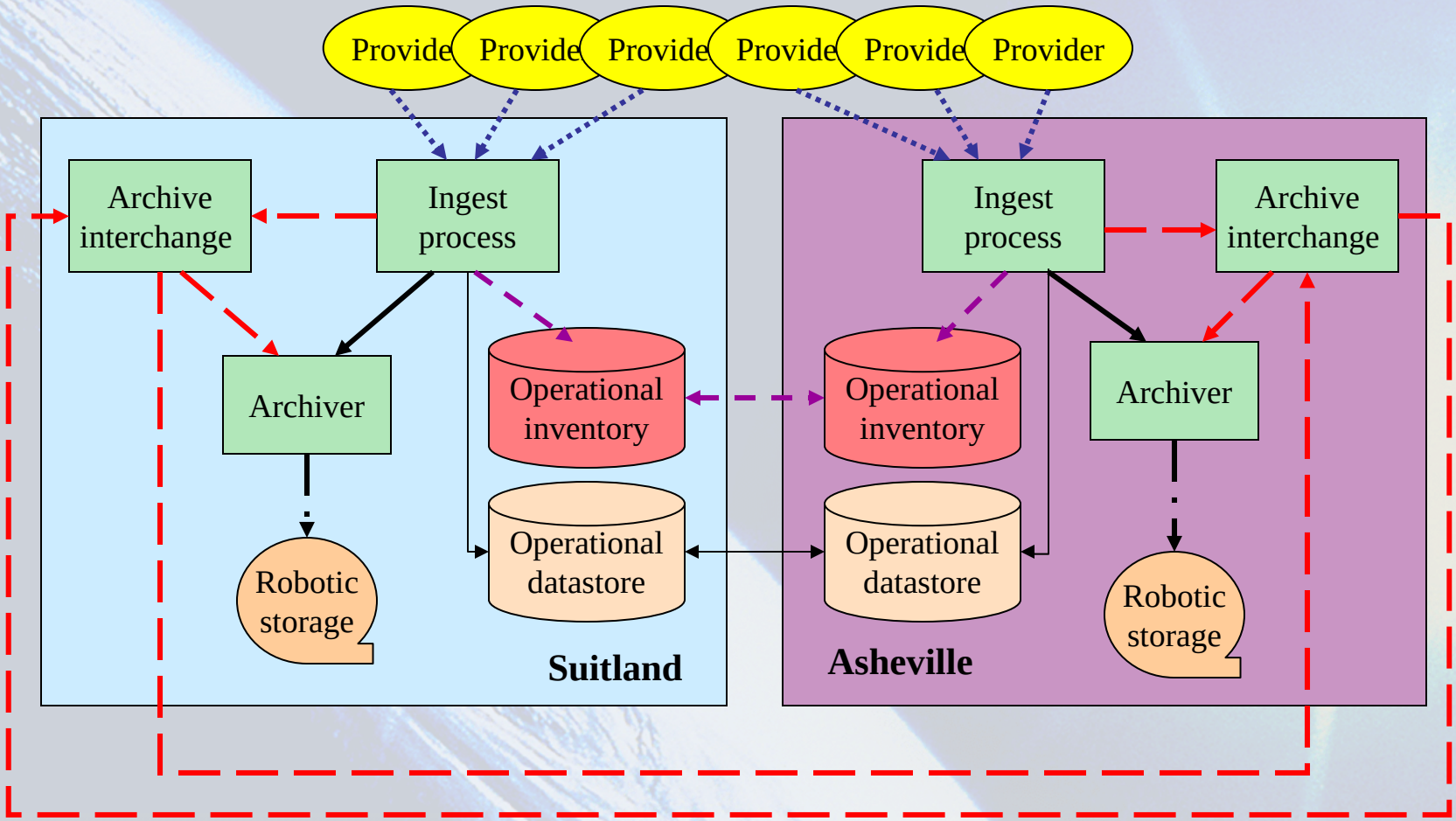
Accomplishments

- Completed CLASS-NPP Requirements Review
- Completed move plan for NSOF, Boulder, and Fairmont
  - Completed Hardware and Network Upgrade Study
- Attained SEI/CMMI Certification Level 2
- Software releases
  - Release 3.3 became operational on June 23, 2005
  - Release 3.3.1 became operational on July 11, 2005
  - Release 3.3.2 is undergoing integration and test
  - Release 3.4 is under development
- Interfaced with NOAA Metadata Manager's Repository (NMMR)
- Certification and Accreditation
  - Working with Mitretek and OSD ITSSO to complete CLASS C&A Package
  - Completed installation of firewalls for all CLASS systems

System Overview



System Overview – Distributed Redundant Archive



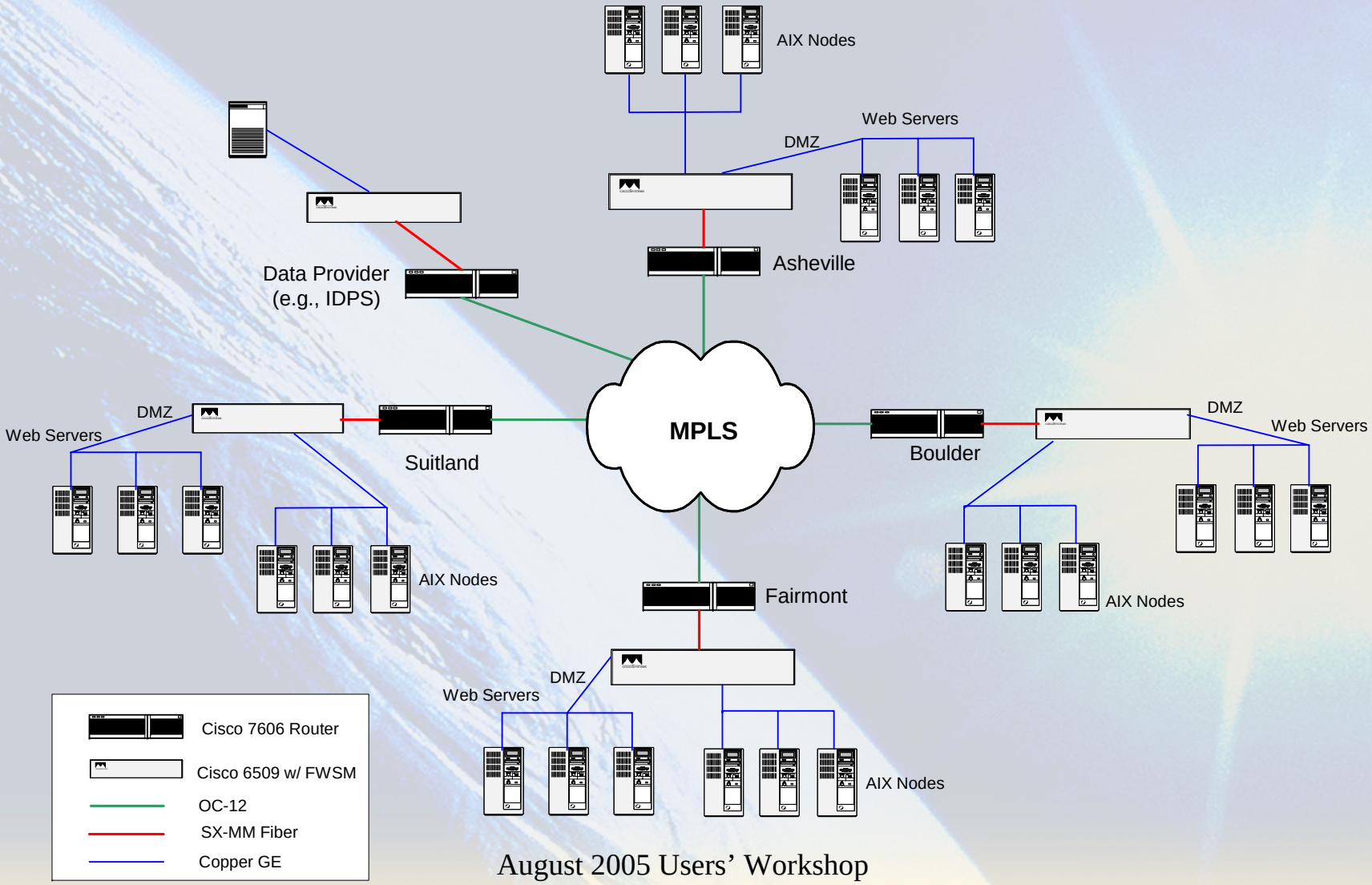
Hardware and Network Design

- CLASS will have three operational sites (NGDC, NCDC, and NASA IV&V)
- CLASS will have a centralized development environment and a centralized integration and test environment at NSOF
- CLASS will use a Multi-protocol Label Switching (MPLS)-based peerless IP network used for site to site communication
- New hardware and network architecture
  - Implemented by January 2006
  - Modular, scalable, and redundant
  - Addresses security requirements

Requirements Affecting Hardware

- Ingest performance requirement (NPP)
  - The System shall be capable of ingesting at least the daily volume of IDPS-produced data, on average, per 24-hour period.
  - The System shall be able to ingest 4 TB of NPP data per 24-hour period
  - Implies additional data replication of 4 TB
- Delivery performance requirement
  - The System shall be capable of distributing 24 TB of data per day.
  - The System shall be capable of producing and distributing 24 TB volume of NPP data and products per day.

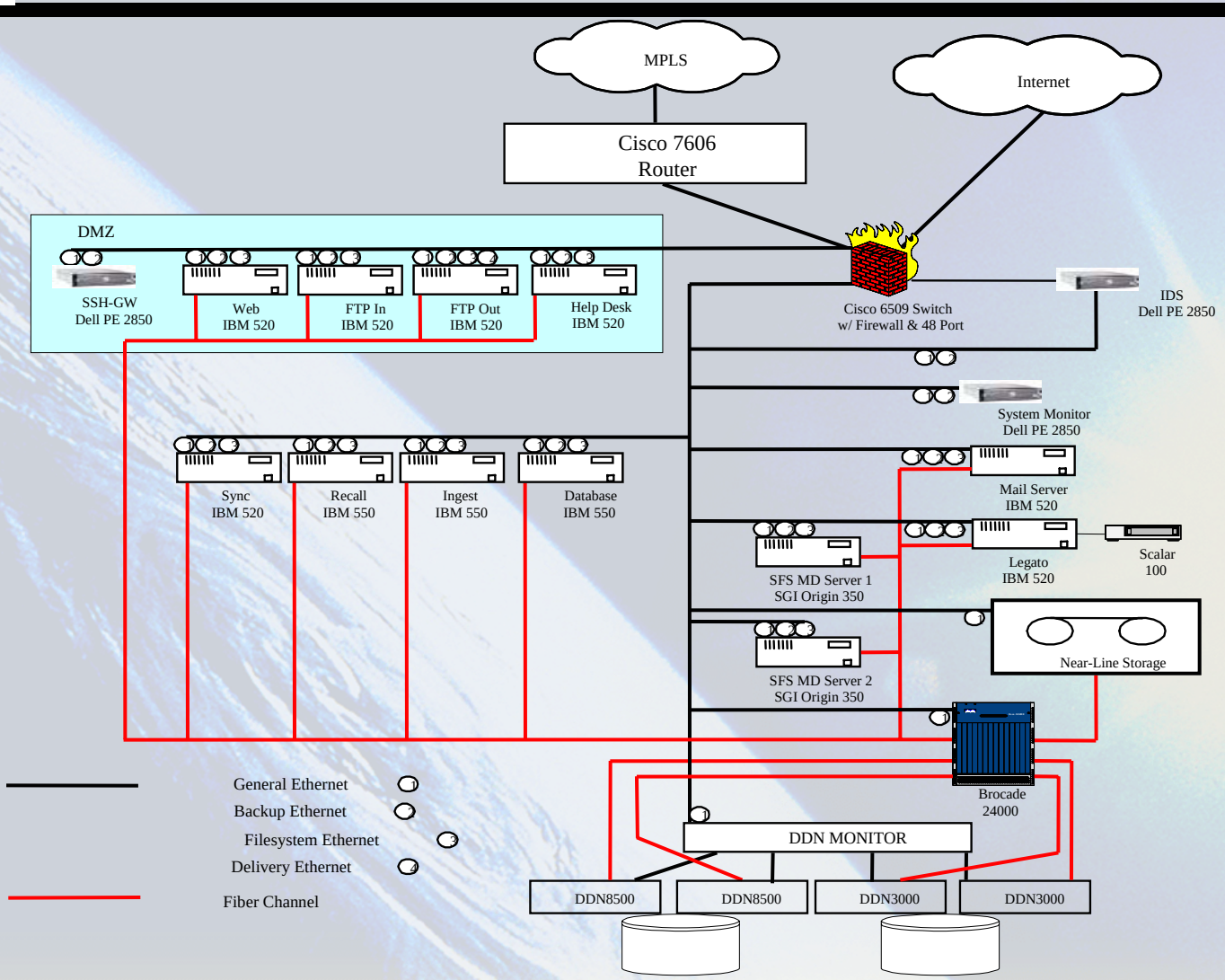
CLASS Multi-Site Configuration



Hardware Design Features

- SAN-based architecture
  - All data transferred via fibre-channel network
- Shared File System (SFS)
  - Eliminates unnecessary data copies
- Private and public security enclaves
  - Public access is only available through DMZ
- System scalability
  - Easy addition of new processors and storage devices
  - Multi-path connectivity

CLASS Single Site Configuration



Hardware – Performance Characteristics (1)

- Fibre-channel SAN bandwidth
  - 2 Gb/sec
  - 5% overhead
- Processors
  - IBM 550
    - 4 - 1.65 GHz Power5 CPUs
    - 8 GB Memory
    - Processor to memory bandwidth – 25.5 GB/sec
    - Remote I/O bandwidth – 8.8 GB/sec
  - IBM 520
    - 2 - 1.65 GHz Power5 CPUs
    - 4 GB Memory
    - Processor to memory bandwidth – 12.8 GB/sec
    - Remote I/O bandwidth – 4.4 GB/sec

Hardware – Performance Characteristics (2)

- SAN storage
  - DataDirect Networks 3000
    - 11 TB of fibre channel (FC) disk
    - 650-700 MB/sec sustained throughput
  - DataDirect Networks 8500
    - 156 TB (usable) of SATA disk
    - 1.5 GB/sec sustained throughput
- Shared File System
  - ADIC StorNext File System
    - Write operations – 63 MB/sec (1 thread)
    - Read operations – 45 MB/sec (1 thread average)
  - Shared File System Metadata Server
    - SGI Origin 350
      - 4 - 700 MHz CPUs
      - 4 GB Memory
      - Processor to memory bandwidth – 3.2 GB/sec
      - Sustained I/O – 1.07 GB/sec

Thank-You

August 2005 Users' Workshop

Background Information

August 2005 Users' Workshop

System Functions – Ingest and Store Data

- Receive and ingest satellite data and derived product data
- Create inventory entries in database
- Create browse image data files for selected data sets; store on-line
- Create netCDF files from selected product data
- Store some files in permanent cache, others in temporary cache
- Archive all files in robotic system in native format

System Functions – Interface with Users

- Log in or set up user profile
- Initiate catalog search
- View search results: catalog data, browse images, dataset coverage maps
- Order data or set up subscriptions
- View order status
- Visualize and download product data netCDF files
- View FGDC-compliant collection-level metadata

System Functions – Access Data

- Invoked by User Interface and Order Processing modules
- Search catalog for data sets that meet user-specified criteria:
  - Data type
  - Time range
  - Geographic coverage
  - Other criteria appropriate for each data type
- Locate files on-line or retrieve files from robotic storage
- After ingest or retrieval, keep files on-line for several days for quick access

System Functions – Visualize Data

- Invoked by User Interface
- Interfaces with tools for creating visualization images
  - Ferret for data analysis
  - Browse Image Generator
  - CoastWatch Image Generator for CoastWatch images
  - McIDAS Image Generator for GOES images
- Returns URL of file containing image to User Interface

System Functions – Process Orders

- Fill orders submitted through User Interface:
  - Put ordered files in FTP area for users to pull
  - Notify users when files are available
- Provide subscription order service:
  - Place orders automatically for newly ingested data sets that meet subscription criteria
  - Push files to subscribers or make files available for pull
- Provide bulk order service:
  - Create orders for large amounts of data that cannot be ordered conveniently through User Interface
  - Retrieve files in time-ordered blocks and place in FTP area for user to pull
  - Notify user when each block is available

System Functions – Maintain, Monitor, Control

- Maintenance
  - Automatically create new log files each day
  - Automatically create new FTP area directories as needed
  - Automatically clean up temporary caches, work directories, FTP area, database tables
- Monitoring and Control
  - Record all activity and error messages in standard log files
  - Send error and warning messages from log files to operators via e-mail
  - Provide tools and operator interface for monitoring and controlling functions
    - View order status; cancel, restart, prioritize orders
    - View ingest statistics; initiate re-ingest
    - View ingest and order activities; stop and restart activities
    - Modify allocation of processes to platforms
    - Modify processing parameters

CLASS Nodes

- CLASS Network Structure (Five Nodes)
  - CLASS-IDPS Point of Presence (PoP), NSOF, Suitland, MD (TBR)
    - IDPS "on-ramp" for NPP data onto the CLASS network
  - CLASS-Asheville, NCDC, Asheville, NC
    - Fully operational CLASS site, functionality includes ingest, distribution, archival, and replication of NPP data
  - CLASS-Boulder, NGDC, Boulder, CO
    - Fully operational CLASS site, functionality includes ingest, distribution, archival, and replication of NPP data
  - CLASS-Fairmont, NASA IV&V, Fairmont, WV
    - Fully operational CLASS site, functionality includes ingest, distribution, archival, and replication of NPP data
  - CLASS-Suitland, NSOF, Suitland, MD
    - Development site; CLASS system software development
    - Test and Integration site; CLASS system testing and integration

System Overview – Subsystem Design (1)

Functions are performed by several subsystems which fall into four (4) design categories

- Data Storage and Distribution Subsystems: *Ingest*, *Subscription Order Generation*, *Data Recall*, *Delivery*
  - Each subsystem consists of several independent processes
  - Data are transferred between processes via the database
  - A workflow engine, the Activity Controller, starts processes as needed, routes work items (data sets, orders, recall requests) to processes in predefined sequences, monitors processes and work item progress
    - Enables simple addition of processes and alteration of processing paths
  - Object-oriented design implemented in C++

System Overview – Subsystem Design (2)

- Servers: *Inventory*, *Visualization*
  - Each server is a resident process designed for quick response
  - Requests and responses are transmitted via socket messages in XML format
  - Object-oriented design implemented in C++

System Overview – Subsystem Design (3)

- User Interface
  - Java/XML-based web interface uses off-the-shelf components:
    - **Apache** – Web Server
    - **Cocoon** - publishing framework
    - **Avalon/Excalibur** - database connection pooling
    - **Tomcat** - servlet engine
    - **LogKit** - message logging
    - Informix Java Database Connectivity (**JDBC**) Driver
  - Java servlets perform special functions
  - Communication with Servers via socket messages in XML format

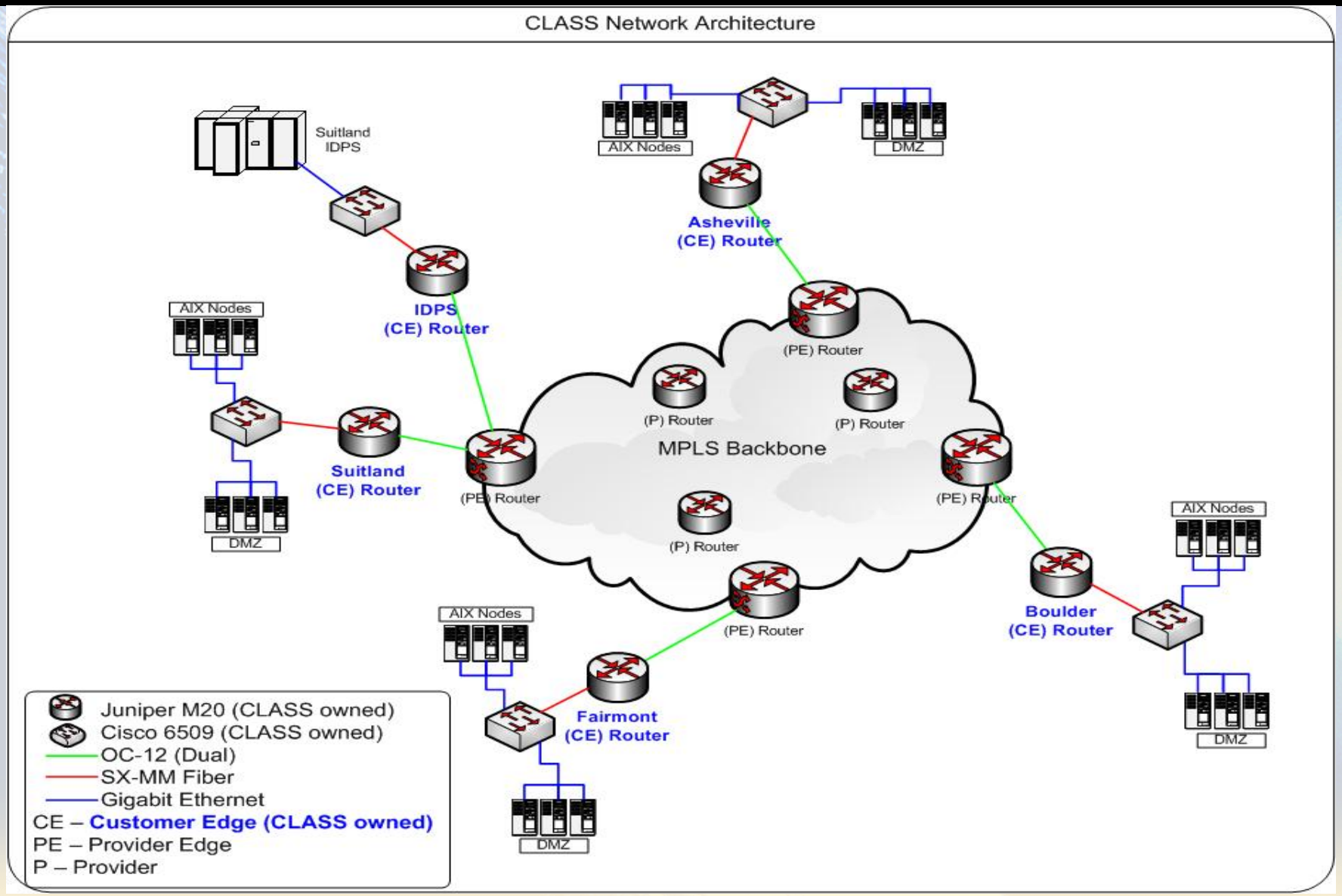
System Overview – Subsystem Design (4)

- Monitoring and Maintenance Tools: *Log Monitor*, *Cache Cleanup*, *Work Space Cleanup*, *Independent Monitoring*, *Operator Interface*
  - Stand-alone tools
  - Various invocation modes:
    - Resident (*Log Monitor*)
    - Run periodically by cron (*Independent Monitoring*)
    - Run as needed by operators or Background Subsystems (*Cache Cleanup*, *Work Space Cleanup*).
  - Various implementations:
    - C++ (*Cache Cleanup*)
    - Perl (*Log Monitor*, *Work Space Cleanup*, *Independent Monitoring*, *Operator Interface*)

MPLS

- MPLS - Multiprotocol Label Switching
  - MPLS is an Internet Engineering Task Force (IETF) initiative that integrates Layer 2 (*Data Link*) information about network links (bandwidth, latency, utilization) into Layer 3 (*Network*) in order to simplify and improve IP-packet exchange. MPLS gives CLASS improved flexibility to divert and route traffic around link failures, congestion, and bottlenecks
  - CLASS will utilize networking philosophies and technologies that are supported by the NOAA Chief Information Officer (CIO), and are the basis for NOAANet
  - MPLS network is currently being implemented by the National Weather Service (NWS) for their NWSnet framework

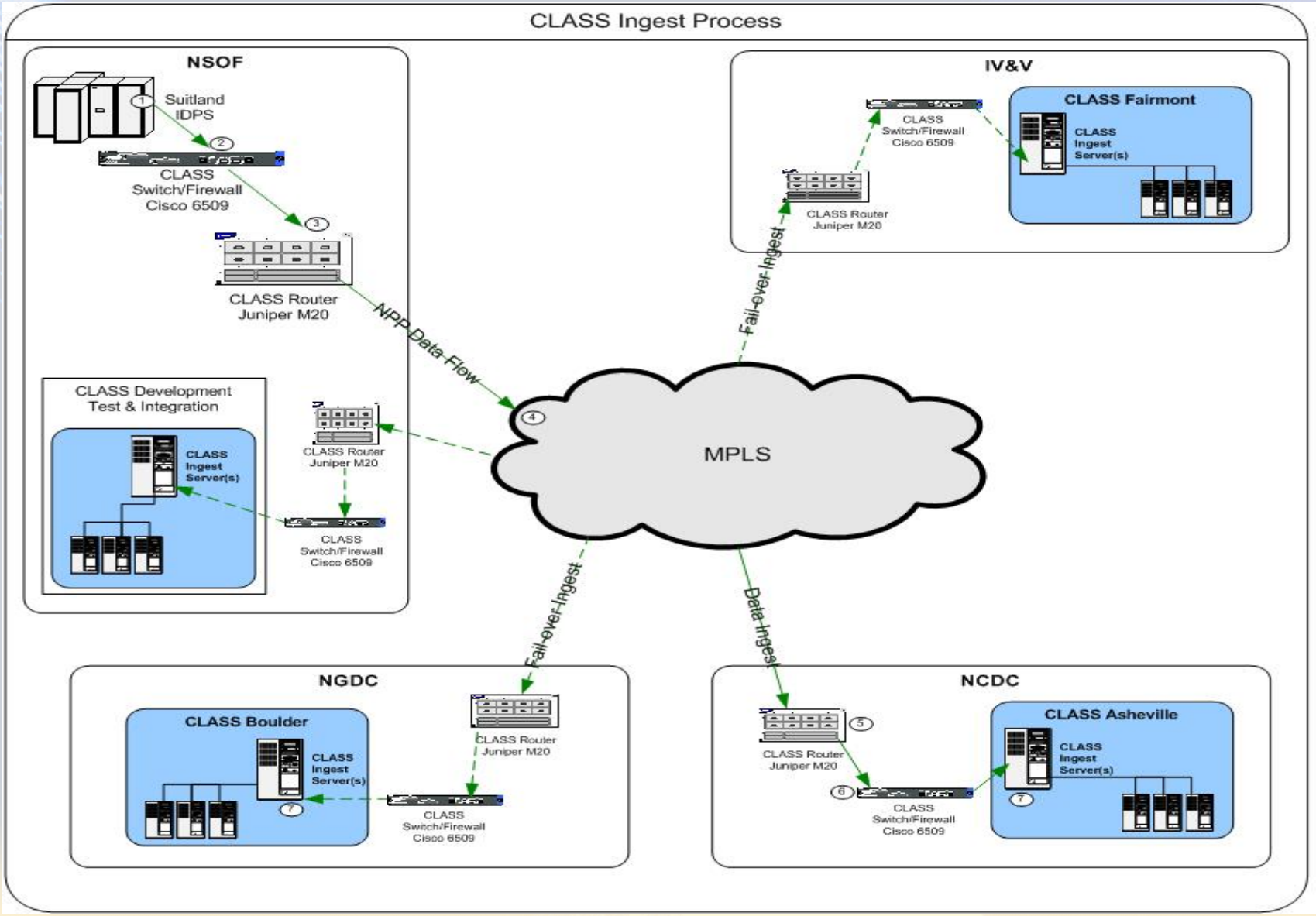
CLASS Network Architecture



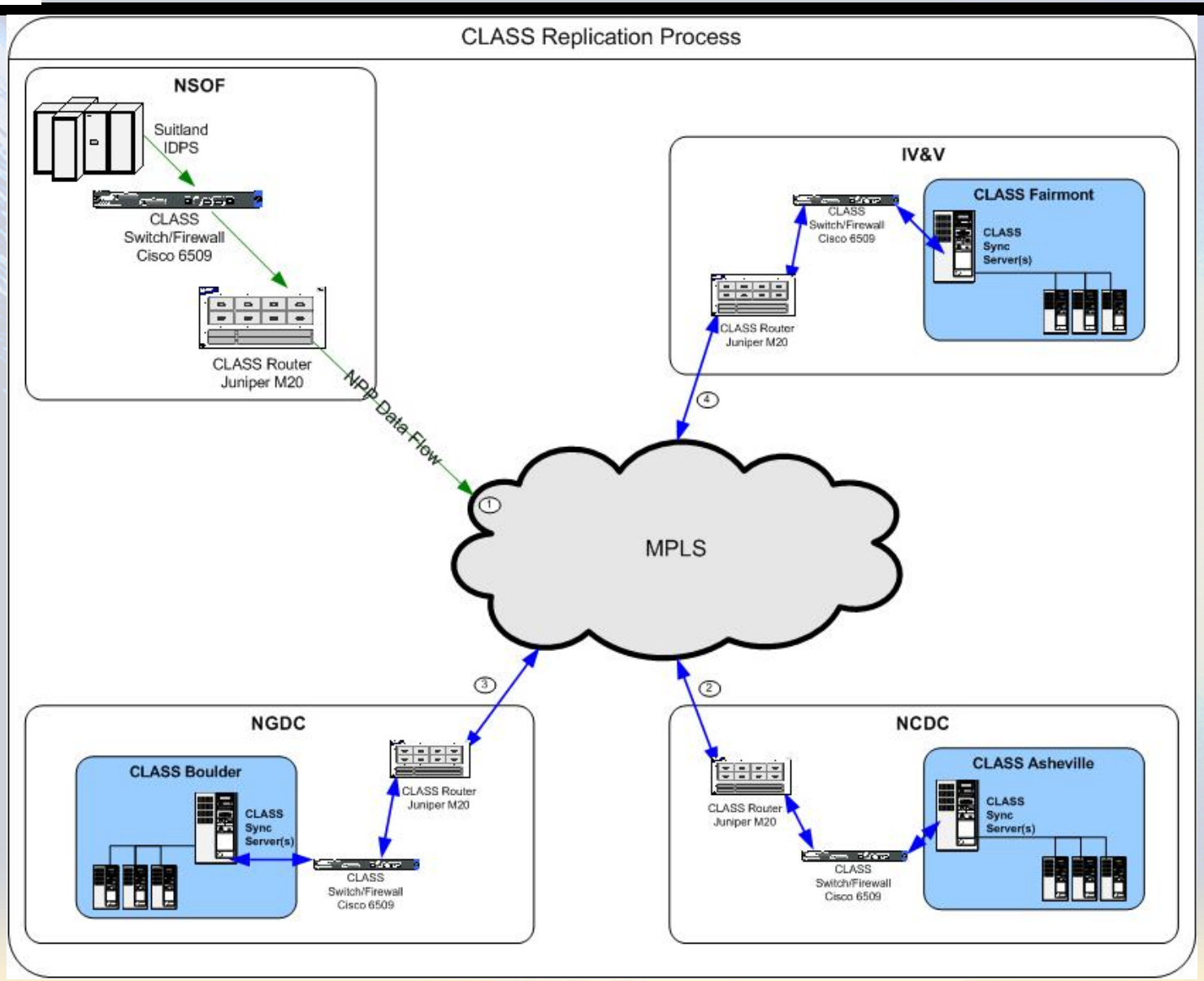
CLASS Network Transport – Ingest and Replication

- CLASS NPP *Ingest* and *Replication* processes will utilize a "Peerless" IP, MPLS network
  - MPLS (Multiprotocol Label Switching) is a versatile protocol that addresses issues of speed, scalability, quality-of-service (QoS), and traffic engineering
  - Meets the bandwidth-management and multi-node service requirements for next-generation Internet protocol-based backbone networks
  - Specifically addresses key CLASS networking requirements
  - "Peerless" IP backbone, completely segmented from the Internet, provides high degree of protection from Denial of Service (DoS) attacks, hacking, and spoofing
  - Cloud topology simplifies the addition of future network nodes and allows for ease in upgradeability for network bandwidth usage

CLASS Ingest Process



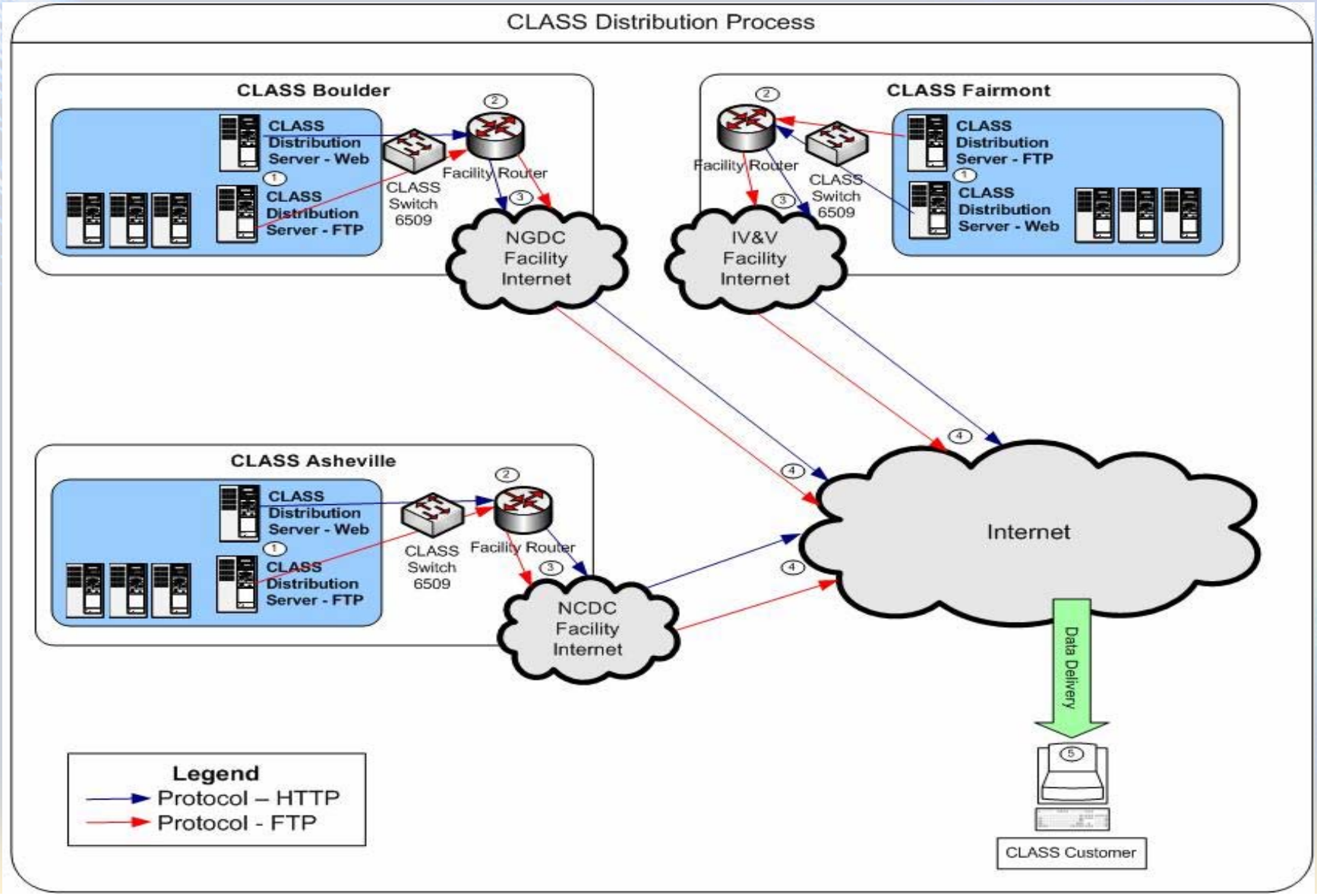
CLASS Replication Process



CLASS Network Transport – Distribution

- CLASS NPP *Distribution* process will utilize the hosting facility's commodity Internet connection to the Internet
- Will not utilize CLASS MPLS network
- Transport protocols:
  - File Transfer Protocol (FTP) push/pull
  - Hypertext Transfer Protocol (HTTP)
- Delivery to NASA's Science Data Segment (SDS) ~2.5 TB/day of NPP data via direct point-to-point circuit (TBR)
- Structural configuration that provides for distribution utilizing load balancing by three operational CLASS sites
  - Asheville
  - Boulder
  - Fairmont

CLASS Distribution Process



Network Hardware

- CLASS Network Equipment and Specifications
  - All CLASS locations consist of identical network hardware build-outs containing redundant equipment to ensure high-availability and network connectivity
  - CLASS hardware will be implemented as “hot-warm” spares to provide redundancy capabilities in the event one should become inoperable
  - CLASS multi-site topology provides for fail-over capabilities, redundancy, and load-balancing characteristics

Network Hardware – Router (1)

- **Juniper M20 Router (TBR) (dual units): the Juniper Networks M20 router will act as the CLASS Customer Edge (CE) router**
  - Fully redundant common hardware components
    - Routing Engines
    - System and Switch Boards (SSB)
    - Fan trays
    - Power supplies
  - Supports 16 ejector-enabled Physical Interface Cards (PICs) via 4 built-in Flexible PIC Concentrators
  - Aggregate Half-Duplex Throughput Performance is 25.6 Gbps
  - Wide Area Network (WAN) uplink up to OC-48 (2,488.32 Mbps data rate)

Network Hardware – Router (2)

- **Cisco 7604 Router (TBR) (dual units): the Cisco 7604 router will act as the CLASS Customer Edge (CE) router delivering high-performance IP/Multiprotocol Label Switching features**
  - Dual supervisor engines and up to two line cards for high availability and redundancy
  - Redundant AC or DC power supplies
  - 7604 supports 4 optical service modules per chassis
  - DS-0 to OC-48/STM-16, Fast Ethernet, Gigabit Ethernet, and 10 Gigabit Ethernet interface cards
  - Backplane bandwidth is 320 Gbps with Supervisor Engine 720
  - Wide Area Network (WAN) uplink up to OC-48 (2,488.32 Mbps data rate)

Network Hardware – Switch

- **Cisco Catalyst 6509, with Firewall Module (dual units)**
  - Features a 40 Gbps interconnection to the Cisco Catalyst 6500 Series Supervisor Engine 720's switch fabric - for high-demand, high-density, Gigabit Ethernet requirements across all 48 ports
  - Provides jumbo frame support, up to 9000 bytes, commonly required in data centers
  - Up to 386 high-performance Gigabit Ethernet over copper ports in a single chassis
  - The Firewall Services Module (FWSM) features multiple contexts (virtual firewalls):
    - 5 Gbps throughput, 100,000 connections per second, and one million concurrent connections
    - Four FWSM can be installed in a single chassis providing scalability to 20 Gbps per chassis
    - Based on Cisco PIX Firewall technology

Performance Characteristics – Ingest and Replication

- Required NPP ingest (inbound) throughput - 4 TB/day = ~388.4 Mbps
- Required NPP replication (outbound) throughput - 4 TB/day = ~388.4 Mbps
- Per site Ingest and Replication volume on CLASS telecom: 776.8 Mbps
- CLASS network support of ingest for NPP data volumes:
  - OC-12 transfer speed is rated at 622.08 Mbps
  - Router - CLASS will utilize two OC-12 interfaces (1244.16) to its MPLS network. This meets bandwidth capability for NPP data volumes
  - Cisco 6509 (with Firewall Module) - 5 Gbps throughput. Four FWSM can be installed in a single chassis providing scalability to 20 Gbps per chassis. This meets bandwidth capability for NPP data volumes
  - MPLS network backbone rated at OC-48 (2,488.3 Mbps), upgradeable to OC-192 (9,952 Mbps). This meets bandwidth capability for NPP data volumes

Performance Characteristics – Distribution

- NPP Data Distribution Volume
  - 24 TB/day = 2330.2 Mbps
  - 2.5 TB/day dedicated to NASA SDS
  - 21.5 TB/day = 2087.4 Mbps
- CLASS network support of distribution for NPP data volumes
  - NPP Data Delivery Bandwidth (Load-balanced - Three Sites)
  - Total delivery: 2087.4 Mbps / 3 = 695.8 Mbps per site
  - Distribution of NPP data will be delivered via the hosting facility's connection to the Internet
  - NASA SDS will receive NPP data distribution via a dedicated circuit capable of transporting 2.5 TB of NPP data daily. (TBR)
- The hosting facilities (NGDC, NCDC, and NASA IV&V) for the CLASS sites have indicated that they are currently, or with network upgrades, able to support the NPP data volume requirements needed to deliver 21.5 TB/day via each facility's commodity Internet infrastructure

Implementation Schedule

- Preliminary Design Status Review – July 2005
- Hardware Upgrade – January 2006
- Network Upgrade – January 2006
- System Design Complete – April 2006
- Operational Sites at NGDC and IV&V – August 2006
- Release 3.4 – October 2005
- Release 4.0 – March 2006
- Release 4.1 – August 2006
- Release 4.2 – December 2006

Implementation Schedule – Release 4.0 Details

- Release 4.0, Operational Date: March 2006
  - Basic NPP Support
    - Ingest by file name
    - Definition of data grouping
    - Database schema changes
    - Basic search and delivery capabilities
    - CLASS - NeS Interface

Implementation Schedule – Release 4.1 Details

- Release 4.1, Operational Date: August 2006
  - Additional NPP Support
    - Ingest HDF5 data
    - Create visualization images
    - Sub-setting of HDF5 files
    - Identify cross-reference information
    - Geographic search capabilities
    - Process Two-Line Element files

Implementation Schedule – Release 4.2 Details

- Release 4.2, Operational Date: December 2006
  - Complete NPP Support
    - Enhancements to web interface for displaying cross-reference information
    - Shopping cart enhancements for support of all NPP delivery options
    - User prioritization
    - Display visualization images in the web
    - Complete subscription and bulk order enhancements
    - Operator interface enhancements including reporting tools
    - Independent monitoring tools

Assumptions

- L2 requirements will be officially approved and will flow into System Specification document
- ICDs will be updated to reflect MRS L2 changes
- SDS does not deliver data to CLASS
- NPOESS Data Exploitation (NDE) does not apply to NPP phase
- Sample HDF5 data will be made available before December 1, 2005
- Data Reconciliation Reports will be sent by the data provider
- Geolocation packaging - Baseline will be preserved
- Software and documentation will come in a bundle from ISF
- IDPS configuration with NOAA will be agreeable to NESDIS Central and CLASS
  - For time delay and aggregation

Issues/Risks

- Undefined and changing requirements make detailed design difficult or impossible
- File naming conventions not completely defined
- Data denial requirements not defined
- Cal/Val requirements not defined
- C3S and ISF Interfaces not fully defined
- Cross-referencing – what does CLASS need to cross-reference
- Availability of sample data
- Data reduction – what effect will this have on data volumes
- Data volume – total volume to be delivered is unknown
- IDPS backup – how will CLASS get data if NESDIS Central is unavailable
- SAN data delivery (no beachhead)
- Quasi-static metadata – unclear how CLASS will obtain and in what format
- Availability of supporting data descriptions – help pages
- Lack of sufficient metadata to support growing user community